# Production Benefits Recognition Equally For Verbs And Nouns

Michael O'Leary

A report submitted as partial requirement for the degree of Bachelor of Psychology with Honours at the University of Tasmania, 2018

## Statement of Sources

I declare that this report is my own original work and that contributions of others have been duly acknowledged.

*Signature*: *Date:* 30/11/2018

## Acknowledgments

There are many people that made this research possible. I would like to thank Michael Garry for his supervision and invaluable assistance during all stages of the process. Jim Sauer for allowing the time for the science to happen. Jan and Paul O’Leary for proofing and for providing, as ever, emotional and material support. Paul Mayne for proofing at a crucial stage. Nick Carlson for providing living conditions that were conducive to science, and Paul Schokman for his assistance navigating statistical packages.

## Table of Contents

## List of Figures

**Figure Title.** **Page.**

Production Benefits Recognition Equally For Verbs And Nouns

Michael O’Leary

Word Count: 8,277

## Abstract

The memory improvement for words that are read aloud compared to silently (the production effect) was examined for people speaking verbs and people speaking nouns. Participants were 3 male and 14 female adults (age range = 19 - 46 years; mean age = 29 years) who spoke English as a first language, with normal or corrected vision. I examined whether the size of the production effect was different for people using verbs compared to nouns. I also examined participants' memory judgements for verbs and nouns following production. There was a large, significant increase in recognition for spoken words compared to silently read words. In contrast, there was no significant difference in the size of the effect when people spoke verbs compared to nouns. A Bayesian analysis indicated weak evidence that people's memory judgments do not differ following production of verbs versus nouns. These results indicate that the production effect is unaffected by word-class, which supports the idea that it could be used to improve teaching and learning outcomes in applied settings.

People tend to have a better memory for words that are read aloud compared to words that are read silently (MacLeod, Gopie, Hourihan, Neary, & Ozubko, 2010). This is known as the production effect. It is a robust effect; the memory benefits are consistently observed in both non-impaired (Pals, Tolboom, Suhre, & van Geert, 2018) and learning impaired populations (Icht, Bergerzon-Biton, & Mama, 2016). And, it is an easy learning strategy that requires less cognitive effort than alternative methods such as elaborating on the information being learned (Bodner & MacLeod, 2016). Thus, its use in applied learning environments, for example to develop more effective learning and teaching strategies, appears to be justified.

However, the research to date has been primarily conducted using nouns and neglecting verbs. Nouns are thought to be easier to remember and are more commonly used in memory research because their use makes investigating underlying processes easier (Kersten & Earles, 2004). But, the exclusive use of nouns in production effect research potentially undermines claims about the real-world memory benefits of production. The aim of this research is, therefore, to address this deficiency in the existing literature by examining the impact of word-class on the size of the production effect.

**Historical Background**

Hopkins and Edwards (1972) are thought to be the first to investigate the memory benefits of production. They proposed that a process occurring during encoding mediates the effect. This was not their primary topic of interest, however, and they did not investigate further. It was another 16 years until the effect was researched again.

Gathercole and Conway’s (1988) research examined the effect of different input modalities (for example, speaking or writing) on long term memory for items, and suggested that the benefit was a result of translating information from one modality to another. For instance, when information presented visually (written words) is translated into an auditory form (speech).

Dodson and Schacter (2001) then conducted research that indicated that the memory benefits of production appeared to be facilitated by the distinctiveness of spoken words. Specifically, they stated that the distinctiveness of the act of speaking study words improved people's memory by reducing people's false memories. However, the effect did not have a formal title at that stage.

Subsequent to that, MacLeod et al. (2010) more specifically defined the production effect. They described it as representing an approximate 10% to 20% increase in memory for spoken words compared to silently read words. Furthermore, they examined the impact of different production and stimulus types on the size of the memory improvement.

While there were gains associated with writing or typing words, the effect appeared to be largest when words were vocalised. Specifically, speaking words at a normal volume was more effective than whispering words, and speaking loudly or singing resulted in a larger improvement than speaking normally (Forrin, MacLeod, & Ozubko, 2012). However, the size of the effect was not as large when non-words were used, or when listening to other people say the words.

Additionally, Forrin, Jonker, and MacLeod (2014) showed that the production effect was unaffected by the simultaneous use of other encoding strategies, such as elaboration. Elaboration refers to the idea that greater activity when encoding an item will lead to greater memory for that item. For instance, the concurrent use of generation (when people generated and tried to remember a word that was similar to a stimulus word) resulted in no difference in the size of the production effect. Instead, the production effect appeared to provide a memory benefit that was additional to other learning strategies. This was taken as evidence for the potential suitability of the production effect in applied settings because it can be used effectively in conjunction with other techniques.

**Mediating Mechanisms**

The research to date indicates that the production effect is mediated by a dual-process mechanism (Fawcett & Ozubko, 2016). This comprises two independent components; distinctiveness and strength.

**Distinctiveness mechanism.**

Broadly, the distinctiveness mechanism is an explicit memory improvement that involves an interaction between the processes used to encode an item and the retrieval strategy that is used to retrieve the information. Essentially, the act of reading a word aloud creates a distinctive production trace compared to silent items. This production trace can then be reproduced in memory tests.

More specifically, there are two factors that are thought to affect the distinctiveness of an item. The first is the quantitative nature of the encoding processes used (Mama & Icht, 2016). There are more processes involved when a person reads an item aloud; for example, auditory processing and motor/articulatory processing (Forrin & MacLeod, 2017). When a study word is silently mouthed instead of spoken, the size of the memory improvement is reduced because mouthing removes a process; hearing the word (Forrin et al., 2012).

Cross-modal translation is another process that is thought to contribute to the distinctiveness mechanism of the production effect (Rackie, Brandt, & Eysenck, 2015). In the case of verbally producing written words, the words must be translated from a visual form into a form that can be spoken. Similarly, when study words are presented in an auditory form there is a production effect after writing the words down. However, when study words are presented and tested in the same form there is no production effect observed. Mama and Icht (2016) have theorised that words that have not been translated lack the associated episodic memories that make them comparatively distinctive. Consistent with Dodson and Schacter's (2001) research, Forrin and MacLeod (2018) have stated that translation benefits memory by reducing false memories.

The second factor that is thought to contribute to an item's distinctiveness is the qualitative nature of the encoding processes that are used. Specifically, Forrin and MacLeod (2017) have stated that uniqueness is the predominant influence on the quality of encoding. In the case of spoken words, the personal relevance associated with hearing oneself speak is the unique aspect. When participants hear other people say the study words the memory improvement is smaller.

Quinlan and Taylor (2013) have provided additional evidence to support the idea that the qualitative nature of the encoding processes used contributes to the memory benefits of production. As previously stated, singing or loudly speaking words results in a larger production effect than speaking words at a normal talking volume. They have reasoned that this difference is a result of the uniqueness and effort involved in singing and speaking loudly compared to speaking normally.

**Strength mechanism.**

The second process that has been argued to mediate the memory benefits of production is the strength mechanism (Fawcett, 2013). Essentially, the act of production strengthens the word's associated representation in memory. This allows the word to be more readily recalled during test conditions. The resulting memory improvement is considered to be an implicit improvement as opposed to an advantage based on an explicit retrieval heuristic.

Fawcett and Ozubko (2016) have reasoned that the strengthened memory trace may be the result of an attentional process. Essentially, participants might pay less attention to non-produced items. They have demonstrated this concept by showing that people are more prone to mind-wandering when reading paragraphs silently compared to reading aloud. However, they note that their theory is largely based on speculation and, if accurate, may only be a partial explanation.

It has been more difficult for researchers find empirical support for the strength mechanism because it accounts for a smaller proportion of memory improvement than the distinctiveness mechanism (Fawcett, 2013). Despite this, it is conceivable that the production effect may more noteworthy as a study aid if the associated memory trace of items is strengthened, rather than wholly relying on distinctiveness. This is because distinctiveness is necessarily relative. While the memory improvement for distinctive items does not come at the expense of less distinctive items, it is conceivable that memory benefits that do not rely on a comparison with other items may more useful in real-world settings (Dewhurst, Rackie, & van Esch, 2016).

**Design Effects**

Empirical support for the dual-process explanation has been derived partly from consistent design effects that occur when investigating the production effect (Fawcett, 2013). The largest and most consistent memory improvements have been found in within-groups designs while small or non-existent improvements are typically observed in between-groups designs.

In within-groups designs participants are given a mixed word list comprising both words to be read aloud and words to be read silently (see MacLeod et al., 2010). Words to be spoken are typically presented in a blue font and words to be read silently are in a white font. After a distractor task, participants are then tested on their memory for the presented words. Participants' memory for words that were spoken is then compared to their memory for words that were read silently. The effect under these conditions is largely attributed to the comparative distinctiveness of words that are spoken aloud.

Ozubko and MacLeod (2010) have used a list discrimination task to provide further empirical support for the distinctiveness account. This involves two groups of participants. In the study phase each group of participants is given two lists. Group A receives a critical list of words to be read both out loud and silently, and an accompanying list of words to be read all aloud. Group B receives a critical list of both aloud and silent words, and an accompanying list of words to be read all silently.

In the test phase participants are presented with words that were in the study phase and asked to identify which list they thought words came from. A production effect for words from the critical list is only observed in Group B, when the accompanying list is all silent as opposed to all aloud. This is because the presence of the all aloud list overrides the distinctiveness of the words spoken out loud off the critical list. Having produced a word does not indicate list status because both lists contain produced items.

Similarly, in between-groups designs the relative distinctiveness of spoken words is removed. In these designs, in the study phase participants are given lists that contain only words to be spoken or only words to be read silently. Their memory for these items is then tested, and the two groups are compared with each other to determine whether a production effect has occurred.

The effect seen in these paradigms has been consistently smaller (or non-existent) than in within-groups designs (Fawcett, 2013). However, that an effect occurs at all is taken to be support for the strength mechanism. This is because there are no distinctive cues available if a list contains only one study method. The items produced are no longer comparatively distinctive and retrieval heuristics that rely on distinctiveness become ineffective.

More recent studies have investigated the dual-process account of the production effect with experimental designs using remember/know judgments (Fawcett & Ozubko, 2016). In the test phase of these designs, participants are asked to distinguish between words that they can explicitly ‘remember’ studying and words that they implicitly ‘know’ were presented in the study phase but lack any associated explicit details. Participants’ ‘remember’ judgments represent the distinctiveness mechanism because access to explicit details is required for a retrieval heuristic to be effective. In contrast, ‘know’ judgements are taken to represent the implicit strength mechanism (Fawcett, 2013).

**Applied Benefits**

The potential applied benefits for the production effect appear to be promising (Bodner & MacLeod, 2016). The 10% to 20% improvement that is consistently observed is substantial (Pals et al., 2018). Moreover, the memory benefits can be seen in various populations. Icht et al. (2016) have observed memory improvements following production in adults who had difficulty producing intelligible speech (dysarthria). Icht and Mama (2015) have observed the effect in five-year-old children, and Lin and MacLeod (2012) have observed the effect in a group of older adults aged 67 to 88 years.

It is also a relatively effortless technique to practically implement compared to other learning strategies such as elaboration that require more cognitive effort (Bodner & MacLeod, 2016). The variety of populations in which the production effect is observed, and the simplicity with which it can be achieved, suggest that the production effect could justifiably be used to improve both learning and teaching strategies in applied settings (Pals et al., 2018).

Furthermore, the production effect has been observed in both recognition (Fawcett, 2013) and recall (Mama & Icht, 2018) tasks. Practically, this suggests that the production effect has the potential to be a suitable learning-tool for real world settings such as multiple-choice tests, as well as tasks that require unaided recall.

However, the effect has been found to be larger in recognition paradigms (Bodner & MacLeod, 2016) This is because the distinctiveness of items becomes less relevant when people do not have words presented to them, so they can decide if they recognise them. Bodner and MacLeod have advised that that care should be taken regarding the type of practical testing applications the production effect is considered for, rather than presuming that it will be equally effective in different settings.

**Verbs Versus Nouns**

Overall, these findings have promising implications for the usefulness of the production effect in applied settings. However, the research to date appears to have been predominantly conducted using nouns and neglecting verbs. The exclusive use of nouns over verbs appears to frequently be the case in memory research more broadly. This is because verbs are thought to be more difficult to remember, thus their use has the potential to confound the results of an investigation into the underlying mechanisms of memory (Kersten & Earles, 2004).

One explanation, with considerable empirical support, for this differential memorability is that there are contextual differences between nouns and verbs (Kersten & Earles, 2004). Nouns represent relatively stable, concrete objects or ideas; for example, 'car' or 'politician'. In contrast, verbs are words that describe actions or states of being; for example, 'run' or 'jump'. Consequently, the meaning of a verb can vary depending on the nouns associated with it. For example, 'run' could be paired with either 'car' or 'politician' and have a different meaning in either instance.

This can be conceptualised in terms of differing semantic organisation. Huttenlocher and Lui (1979) have argued that nouns are organised in a hierarchical structure whereas verbs are organised as matrices. Essentially, nouns may be strongly related to a small group of other words whereas verbs may be weakly associated with a larger group of words. As a result, verbs are less likely to be meaningful and memorable when presented by themselves.

If the size of the production effect is affected by these word-class-based memory differences, then it appears to follow that care should be taken when considering integrating the effect into real-world learning environments. In contrast, if the production effect does generalise to verbs, this would support Bodner and MacLeod's (2016) suggestion that the production effect could be useful as a simple and effective learning technique in applied settings. Furthermore, the recently developed theory of embodied cognition indicates that the production effect may be larger in verbs than nouns.

**Embodied Cognition**

Embodied cognition is a recent theory of human cognition in which our understanding of the world and our states of mind depend on our bodies and the actions we perform (Wilson, 2002). The fundamental principle of embodied cognition is that mind is a product of the brain; brains evolved first to move our bodies and mind evolved later. This means our conscious thought evolved in a body- and action-based context.

Consequently, information that is represented in people's sensory motor systems is thought to comprise conceptual content (Mahon & Caramazza, 2008). For instance, a person's sensory-motor representations may contribute to their self-concept. Essentially, feedback from body position can affect a person's idea about who they are and what is important to them. For example, making a fist has been argued to strengthen the association between self-concept and power. Physical evidence for the conceptual element of embodied cognition has been derived from the observation that motor and sensory systems are activated during conceptual processing (Schubert & Koole, 2009).

Additionally, language may be embodied (Mahon & Caramazza, 2008). In traditional views, language is thought to rely on abstract symbols that are arbitrarily related to the objects or ideas they represent. However, according to an embodied cognition approach, cognitive representations and the operations that are performed on them are based on physical context. For instance, the meaning of the word 'hammer' might be stored as the motor representations of the actions that are associated with the act of hammering.

It follows that knowledge could also be embodied. It could be represented in body-and action-based terms. Lakoff and Johnson (2003) point out that many of the metaphors we use in everyday conversation are action based (e.g., ‘This argument is dragging me down.’ ‘You hit the nail on the head.’). If it is correct that knowledge is action based, then reading verbs aloud may combine action meaning with production. This might provide the context that verbs previously lacked, causing them to become more memorable.

**Hypotheses**

Overall, I expected to observe a production effect. It is a reliable effect that has been observed under many different experimental conditions (MacLeod et al., 2010). Furthermore, I used mixed study lists in this experiment; containing both words to be read aloud and silently. This design has been found to encourage the largest production effect (Fawcett, 2013). Thus, my first hypothesis was that recognition for words studied aloud would be significantly better than for words studied silently, irrespective of word-class.

Additionally, I expected production to benefit verbs more than nouns. Based on the difference in memorability between word-classes discussed by Kersten and Earles (2004), I reasoned that the proportion of silently read verbs that were recognised would be smaller than the proportion of silently read nouns. If production interacted with embodied cognition to result in an equal proportion of spoken verbs and spoken nouns being recognised, this would result in a larger production effect for verbs. Therefore, my second hypothesis was that the production effect would be larger in the group using verbs than the group using nouns.

Lastly, I expected that the personal nature of the action meaning provided by embodied cognition (Lakoff & Johnson, 2003) would be most likely to affect the qualitative aspect of the distinctiveness mechanism. Thus, my third hypothesis was that if a difference in the size of the production effect between verbs and nouns was observed, then it would be reflected in a greater increase in ‘know’ judgments for spoken verbs than spoken nouns.

# Method

## Participants

Participants were 3 male and 15 female adults (age range = 19 - 47 years; mean age = 30 years) who spoke English as a first language, with normal or corrected vision. One participant's data was not usable due to an apparent misunderstanding of the instructions (during debriefing they verbally reported recognising several study words but failed to indicate that during the task). Therefore, data from 3 male and 14 female adults (age range = 19 - 46 years; mean age = 29 years) was ultimately used in the analysis.

Participants were partly recruited from the University of Tasmania first-year and third-year psychology student population on the Sandy Bay and Cradle Coast campuses. These students participated for course credit. Raffles of supermarket gift vouchers were also used to encourage non-psychology students to participate. Advertisements were posted in public areas around the Sandy Bay campus including on bulletin boards, on the Division of Psychology's online research participation system (SONA), and in a third-year cognitive neuroscience unit.

When participants signed the consent form for the study, they were assigned a unique four-character code comprising two random letters and two random digits. This code was recorded on their consent form, and in an electronic master participant file, and was the only identifier used in all electronic and hardcopy copy forms and files.

A *t*-test approach in G Power 3.0 (Faul, Erdfelder, Lang, & Buchner, 2007) was used to perform power calculations to determine the approximate number of participants required to observe a difference in the size of the production effect between verbs and nouns. Kersten and Earles (2004) found a large increase in memory from silently read verbs to silently read nouns (Cohen's $d > 0.80$). Given that this difference in memorability is the basis for the expected difference in the size of the production effect, an effect size similar to Kersten and

Earles's figure was considered possible. Thus, a relatively conservative moderate effect size (Cohen's $d = 0.65$) for the difference in the size of the production effect was estimated. Therefore, for 80% power with $\alpha = .05$, at least 78 participants were calculated to be required.

However, it was recognised that it was unlikely to be possible to recruit 78 participants within the available time-frame. Instead, it was decided to attempt to recruit as many people as possible, and to consider whether the direction of any effect was in the predicted direction when interpreting the results. To further examine whether a differential production effect based on word-class was observed, a Bayesian analysis of the differences between was also used.

Compounding the effect of the time-constraints of an Honours project on participant numbers, the ethics application for this study was submitted, and therefore approved, later than planned. These restrictions resulted in an exploratory and descriptive design, essentially a pilot study involving 17 participants.

**Apparatus/Instrumentation/Materials**

PsychoPy 1.90.1 software on an Asus F540L laptop with a 15.6-inch colour monitor was used to run the experiment. Consistent with several studies that have examined the production effect (Bodner & Taikh, 2012; Jones & Pyc, 2014; MacLeod et al., 2010), the item pool consisted of two lists of 100 words between five and ten letters long. One list comprised verbs and the other list comprised nouns.

From these lists, 80 verbs and 80 nouns were used in the study phase (see Appendix A for item-pools). To control for order effects, four randomly ordered study lists for the 80 verbs and 80 nouns were created. While the words were presented in a different order in each list, the same reading method (aloud or silent) was used for each individual word.

For the test phase, four randomly-ordered test lists for nouns and four randomly-ordered test lists for verbs were created. Each test list comprised 60 words. Forty of those words (20 aloud and 20 silent) were on the study list and 20 words were previously unseen foil items. These lists contained the same words however the order of item presentation was varied.

**Procedure**

The procedure used was based on the method most commonly adopted when investigating the production effect (see MacLeod et al., 2010; Fawcett & Ozubko, 2016). To control for auditory interference, participants were tested separately in a single session which took approximately 30 minutes to complete. After reading the information sheet, participants were told that there would be three parts to the experiment; a study phase, a filler task, and a test phase. The study phase was described, and they were given the opportunity to ask any questions they had about what they were being asked to do. To ensure intentional learning was tested, participants were told that they were learning words for a later memory test. No further information regarding the details of the test phase was provided to participants at this stage.

Consistent with Quinlan and Taylor (2013), the first part of the study phase was a familiarisation phase to accustom participants to the colour coding for different instructions. This comprised four practice words (adjectives) displayed consecutively in the centre of the computer screen in letters three centimetres high against a black background. Two of the words were in a blue font, indicating that they were to be read aloud. Two of the words were in a white font, indicating that they were to be read silently.

During the main part of the study phase, each participant was presented with one of the 80-word study lists. Forty of these words were in blue and 40 words were in white. Items were presented for 2000-ms followed by a 500-ms blank screen before the next word was presented for 2000-ms followed by a 500-ms blank screen and so on until all words had been displayed.

Between the study and test phase, a backwards-counting filler task (Han & Kim, 2004) was used to ensure that participants were not retaining the study words in their working memory. Participants were asked to choose a number between 100 and 110 then count backward in groups of threes from that number until they went below 50. This took approximately 60-120 seconds.

The test phase, comprised of a recognition task, was completed immediately following the filler task. The instructions were explained to participants and they were asked if they had any questions. Participants were then requested to explain back to the researcher what they understood the instructions to be. Any misunderstandings were clarified prior to the test phase commencement.

During the recognition task, participants were presented with a test list consisting of 60 words. The words were consecutively displayed in yellow text so there were no overlaps with the colour of previously presented words. Participants were asked to discriminate between ‘old’ words (slash key – right hand) that were on the study list and ‘new’ words (Z key – left hand) that were not. Timing for the delivery of items in this phase was determined by participants’ key presses; once a word was presented it stayed on the screen until the participant responded, only then was the next word presented on the screen.

If a participant responded that they recognised a test word from the study list, then they were then asked to make a familiar/know judgement. A caveat regarding these judgments is that they are most reflective of the underlying mechanisms of memory when clear instructions are given (Fawcett & Ozubko, 2016). Therefore, to facilitate ease of understanding, the terminology that was used was changed from the ‘remember/know’ terms that are frequently used to ‘know/familiar’.

For these judgements, if the participant felt that they had access to explicit details, such as whether they read it aloud or what they thought when they read the word, then they were asked to respond that they ‘know’ the word (slash key – right hand). Alternatively, if the participant felt that they implicitly recognised a word but lacked any associated explicit details, then they were asked to respond that the word was merely ‘familiar’ (Z key – left hand).

**Design and Analyses**

The analyses comprised two parts. The first analysis was designed to answer hypothesis one and two, the second analysis addressed hypothesis three.

**First analysis.**

For the first analysis, the first independent variable was ‘reading-method’. This was within-groups, with two levels. All participants were given study lists comprising both words to be spoken and words to be read silently. This variable was used to test for the production effect.

The second independent variable was ‘word-class’. This was between-groups, with two levels. Participants were randomly assigned to receive either study and test lists containing verbs or study and test lists containing nouns. This variable was used to test for a difference in the size of the production effect when using verbs compared to using nouns.

The dependent variable was ‘item-recognition’. This was a continuous variable measured as the proportion of words from the study list that participants correctly recognised (hit-rate). Data analysis on these variables was conducted using a 2X2 mixed ANOVA with SPSS version 23. The difference in recognition between spoken and silently read words was quantified using 95% confidence intervals and standardised effect sizes (Cohen’s *d*).

Following the ANOVAs, Bayesian analyses were conducted to calculate Bayes factors to quantify the strength of evidence in favour of the experimental hypothesis ($BF_{10)}$. This was done using the *t*-test methods in Jamovi 0.9.5.8-win64 with a default prior. A Bayes factor greater than 3.0 would be considered moderate evidence in favour of my experimental hypothesis, while a Bayes factor less than 0.33 would be considered evidence in favour of the null hypothesis (Wagenmakers et al., 2018).

One Bayes factor was calculated to examine whether a production effect was observed, overall. This involved comparing participants' recognition for spoken words compared to silently read words, irrespective of word-class.

A second Bayes factor was calculated to examine the belief that verbs are more difficult to remember than nouns when silently read. Because the reduced memory for verbs was the basis for the hypothesised difference in the size of the production effect between word-classes, the hit-rate for silently read verbs was compared to the hit-rate for silently read nouns.

A third Bayesian calculation was used to examine the belief that the memory benefits of production would be greater when using verbs than when using nouns. This involved comparing the size of the difference in the proportion of recognised study-items between silent and aloud words in the group that used verbs to the group the used nouns.

D-prime (d') measures of sensitivity were also calculated to determine the probability that the proportion of recognised study words occurred under the null hypothesis of chance performance. Higher values indicate better sensitivity. Essentially, this was used to quantify the extent to which people were able to discriminate between study words and foil items. The overall d' was calculated using the total proportion of hits and false alarms. Also, to have a greater understanding of what the overall d' represented, the d' was calculated separately using only aloud words and false alarms and only silent words and false alarms.

Lastly, the d' was calculated separately for both verb and noun groups as an indication of whether word-class affected the ability of participants to discriminate between test items. An independent samples *t*-test was used to examine whether there was a difference in discrimination in the group that used verbs compared to the group that used nouns.

**Second analysis.**

The second analysis, aimed at investigating hypothesis three, was used to identify the mechanism that was being manipulated by production. This involved examining the dependent variable 'item-recognition' in terms of the proportion of 'know' and 'familiar' judgments that participants made for words that were read aloud compared to silently.

Six paired-samples *t*-tests were conducted using SPSS to quantify and analyse the difference between the memory judgments that participants made when reading silently compared to reading aloud. The first four tests consisted of comparing 'know' to 'know' judgements and 'familiar' to 'familiar' judgments for aloud and silent words. The sizes of the differences were then compared between the verb group and the noun group using 95% confidence intervals and standardised effect sizes (Cohen's *d*). The Cohen's *d* value was calculated for a within-subjects effect to correct for the correlation between conditions (Dunlap, Cortina, Vaslow, & Burke, 1996).

Additionally, a Bayesian analysis was conducted to examine the hypothesis that production would increase participants' 'know' judgments (distinctiveness mechanism) more in the verb group than the noun group. This involved using a default prior and comparing the size of the difference between the proportion of 'know' judgments for silently read words and spoken words, in the verb group versus the noun group.

A final Bayesian analysis was conducted to examine the belief that production would not increase participants' 'familiar' judgements (strength mechanism) more in the verb group than the noun group. This involved using a default prior and comparing the size of the difference between the proportion of 'familiar' judgments for silently read words and spoken words, in the verb group versus the noun group.

The proportions that were used for the second analysis were calculated using the independent 'know/familiar' method (Fawcett & Ozubko, 2016). Using this method, the proportion for ‘know’ judgments was calculated as the proportion of study words presented in the test phase that were correctly identified. In contrast, the proportion for ‘familiar’ responses was adjusted. This is because the processes that mediate the explicit and implicit memory benefits of production operate independently of each other, as opposed to being mutually exclusive. Thus, a produced study word may receive both an implicit and explicit memory improvement.

However, when using the ‘know/familiar’ paradigm, participants are asked to respond ‘familiar’ only if they have no explicit details associated with an item. In contrast, ‘know’ responses do not have a similar restriction on them. It would not be logical to ask participants to respond ‘know’ only when they can recall explicit details about a word but have no implicit feeling that they had seen that word before.

Therefore, unadjusted proportions make the assumption that if an explicit benefit occurred then an implicit benefit did not occur (Fawcett & Ozubko, 2016). This would potentially underestimate the number of study words associated with an implicit memory improvement. To adjust for this bias, ‘familiar’ responses were measured as the proportion of study words that remained after the words that participants had responded ‘know’ to were removed. By including a reference to the number of opportunities a person had to say ‘familiar’ a more accurate measure was gained. Practically, this meant that if there were few opportunities to respond ‘familiar’ once the ‘know’ responses had been accounted for, then a person’s 'familiar’ score was adjusted upwards.

To examine the effect this adjustment had on the analysis, the last two paired-samples *t*-tests compared the difference in the unadjusted proportion of participants’ ‘familiar’ responses between silently read words and spoken words. The sizes of the differences were then compared between the verb group and the noun group using 95% confidence intervals and standardised effect sizes (Cohen’s *d*).

# Results

## First Analysis

The first analysis comprised a 2X2 mixed ANOVA. Levene's test showed that for the proportion of recognised study words that were spoken, the variances did not differ significantly between verb and noun groups, $F(1, 15) = 0.50$, $p = .492$. In addition, for the proportion of recognised study words that were read silently, the variances did not differ significantly between verb and noun groups, $F(1, 15) = 1.07$, $p = .318$. Together these results indicate that the homogeneity of variances assumption was met.

There was a large, significant main effect for reading-method, $F(1, 15) = 24.27$, $p < .001$, Cohen's $d = 1.28$. Overall, participants ($n = 17$) recognised a greater proportion of spoken study words ($M = .81$, 95% CI [.72, .90], $SD = .17$) than silently read study words ($M = .61$, 95% CI [.54, .69], $SD = .14$). This represents a 20% (.20) improvement in memory for spoken words compared with silently read words, consistent with the production effect. Figure 1 displays the increase in recognition from silently read to spoken words.

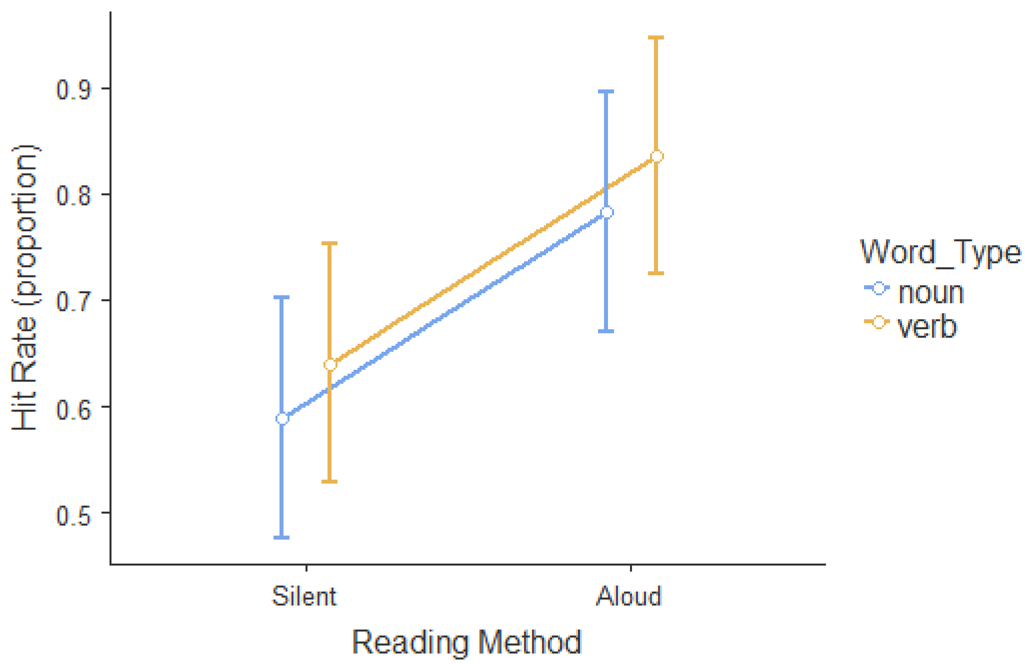


*Figure 1*. Hit-rates for verbs and nouns when speaking and reading silently. *Note*: Error bars represent 95% confidence intervals.

The Bayes factor ($BF_{10}$) for the difference in hit-rate between silently read words and spoken words was 266. This represents extreme evidence in favour of the experimental hypothesis ($BF_{10} > 3$) that words that are spoken are more memorable than silently read words.

In contrast, there was no statistically significant main effect for word-class, $F(1, 15) = 0.61$, $p = .446$, Cohen's $d = 0.52$. Overall, there was no statistically significant difference in the proportion of study words recognised between participants in the noun group ($n = 8$, $M = .68$, 95% CI [.58, .79], $SD = .11$) and participants in the verb group ($n = 9$, $M = .74$, 95% CI [.64, .83], $SD = .12$).

There was no statistically significant interaction between word-class and reading-method $F(1, 15) < 0.01$ , $p = .993$. In the verb group ($n = 9$), participants recognised a proportion of .19 (19%) more spoken study words ($M = .83$, 95% CI [.71, .95], $SD = .14$) than silently read study words ($M = .64$, 95% CI [.54, .74], $SD = .16$) (Cohen's $d = 1.26$). This was approximately equal to the size of the improvement in recognition (.19; Cohen's $d = 1.13$) in the noun group ($n = 8$) from silently read study words ($M = .59$, 95% CI [.48, .70], $SD = .13$) to spoken study words ($M = .78$, 95% CI [.65, .91], $SD = .20$).

The Bayes factor ($BF_{10}$) for the difference in hit-rate between silently read verbs and silently read nouns was 0.28. This indicates moderate evidence in favour of the null hypothesis ($BF_{10} < 0.33$) that there are no memory differences between word-classes.

The Bayes factor ($BF_{10}$) for the difference between the proportion of recognised spoken words and silently read words in the verb group versus the noun group was 0.42. This represents anecdotal evidence in favour of the null hypothesis that the production effect is equal for nouns and verbs.

Using the overall hit-rate proportion of recognised study words (.71) and the overall false-alarm rate (.26), d' = 1.20. Participants' ability to discriminate between study and foil items was 1.20 standardised units better than chance performance. D-prime for the hit-rate of silently read study words (.61) and the overall false-alarm rate (.26) was 0.92. Finally, the d-prime for the hit-rate of spoken study words (.81) and the overall false-alarm rate (.26) was 1.52.

Within the noun group, using the proportion of correctly identified study words (.69) and the false-alarm rate (.24); d' = 1.20. Participants' ability to discriminate between study and foil items was 1.20 standardised units better than chance performance. Within the verb group, using the proportion of correctly identified study words (.74) and the false-alarm rate (.27); d' = 1.26. Participants' ability to discriminate between study and foil items was 1.26 standardised units better than chance performance.

An independent samples *t*-test was conducted to examine whether discrimination was affected by word-class. Levene's test showed that the variances did not differ significantly between verb and noun groups, $F(15) = 0.17$, $p = .683$. These results indicate that the homogeneity of variances assumption was met. There was no statistically significant difference in participants ability to discriminate between study and foil items in the group that used verbs ($n = 9$) compared to the group that used nouns ($n = 8$), $t(15) = 0.26$, $p = .799$.

**Second Analysis**

The second analysis comprised six paired-samples *t*-tests. Within the verb group, the proportion of participants' ($n = 9$) 'know' judgements for silently read verbs was $M = .32$, $SD = .11$. The proportion of 'know' judgments for spoken verbs was $M = .56$, $SD = .23$. There was a large, significant increase ($M = .24$, 95% $CI_{diff}$ [.12, .35], $SD = .15$) from silent to spoken verbs $t(8) = 4.79$, $p = .001$ , Cohen's $d = 1.33$.

Within the noun group, the proportion of participants' ($n = 8$) 'know' judgements for silently read nouns was $M = .29$, $SD = .17$. The proportion of 'know' judgments for spoken nouns was *M* = .47, *SD* = .30. There was a moderate, significant increase (*M* = .18, 95% $CI_{diff}$ [.02, .35], *SD* = .20) from silent to spoken nouns *t*(7) = 2.61, *p* = .035, Cohen's *d* = 0.74.

The Bayes factor ($BF_{10}$) for the difference in the proportion of participants' 'know' judgments between silently read words and spoken words, in the verb group versus the noun group was 0.70. This indicates anecdotal evidence in favour of the null hypothesis that there is not a differential effect of production on participants' 'know' judgments for verbs and nouns.

Within the verb group, the proportion of participants' (*n* = 9) 'familiar' judgments for silently read verbs was *M* = .47, *SD* = .23. The proportion of 'familiar' judgements for spoken verbs was *M* = .59, *SD* = .33. There was small, non-significant increase (*M* = .12, 95% $CI_{diff}$ [.08, .33], *SD* = .27) from silent to spoken verbs *t*(8) = 1.37, *p* = .208, Cohen's *d* = 0.42.

Within the noun group, the proportion of participants' (*n* = 8) 'familiar' judgments for silently read nouns was *M* = .38, *SD* = .25. The proportion of 'familiar' judgements for spoken nouns was *M* = .49, *SD* = .34. There was a small, non-significant increase (*M* = .12, 95% $CI_{diff}$ [-.06, .29], *SD* = .21) from silent to spoken nouns *t*(7) = 1.61, *p* = .151, Cohen's *d* = 0.37.

The Bayes factor for the difference in the proportion of participants' 'familiar' judgments between silently read words and spoken words, in the verb group versus the noun group, was 0.42. This indicates anecdotal evidence in favour of the null hypothesis that there is not a differential effect of production on participants' 'familiar' judgments for verbs and nouns. Figure 2 displays the proportions of 'know' and 'familiar' memory judgements following production of verbs and nouns.

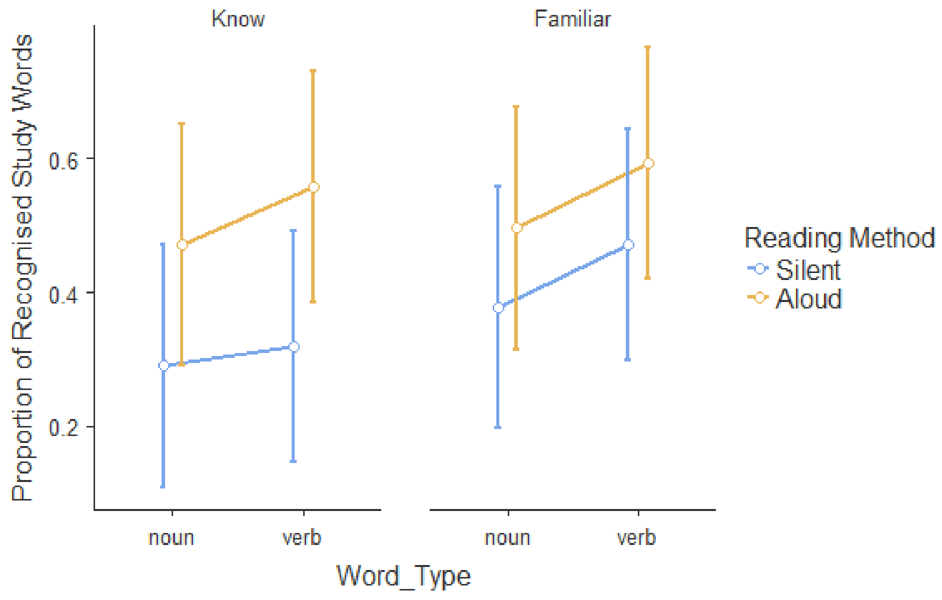


*Figure 2*. Difference in proportions of 'know' and 'familiar' memory judgements following production of verbs and nouns. *Note*: Error bars represent 95% confidence intervals.

## Discussion

Whereas my results are in clear agreement with prior research on the production effect, the hypothesis that word-class would affect the size of the production effect was not supported. Additionally, my results do not support the hypothesis that the distinctiveness mechanism would be differentially affected by word-class. The results of the d-prime calculations justify the analysis of hit-rates because participants were sensitive to study list words. Moreover, the observed level of sensitivity did not differ significantly between verb and noun groups.

### Hypothesis One

The results support my hypothesis that a production effect would be observed. It appears that people are better able to recognise spoken words than silently read words. Consistent with the literature (MacLeod et al., 2010) the size of this effect is large, representing an approximate 20% increase in the proportion of total words that are recognised. Furthermore, the Bayes factor represented extreme evidence that the belief in a production effect is justified.

That the production effect is substantial in size and reliably observed are proposed to be amongst the predominant reasons why the production effect could be successfully integrated into real-world learning settings (Bodner & MacLeod, 2016). The large effect that I observed is consistent with this claim.

**Hypothesis Two**

The results do not support my hypothesis that the memory benefits of production are greater when people speak verbs compared to nouns. The magnitude of the effect for both verbs and nouns is large (Cohen's $d$ = 1.26 and 1.13). This represents an approximate 19% increase in the proportion of total words recognised, irrespective of word-class.

That I didn't observe a greater size production effect for verbs appears to be logical. The hypothesised difference in size was based on the idea that there are contextual differences between word-classes that reduce memorability for silently read verbs (Kersten & Earles, 2004). However, for the words I used, there do not appear to be any inherent differences in memorability between word-classes when silently reading. The Bayesian factor represents moderate evidence for the null hypothesis that there is no difference in memorability between the words that I chose to represent each word-class. Thus, there is not the hypothesised opportunity for a greater size memory improvement for verbs than nouns when reading aloud.

Given these findings, it may be an over-simplification to focus entirely on broad categories such as word-class in the context of memory research. To claim that an effect seen for one word in that class should be seen for all words may fail to account for the varied meanings that words might have across individuals. For instance, there were differing voluntary reports made by participants during debriefing about which words were particularly relevant for them.

For example, one participant who read the word ‘relax’ silently found it instructive. They had not realised that they had been sitting tensely and took it as an instruction. That was memorable for them, however it is conceivable that other less anxious people might not have found the word as memorable. Test anxiety affects people differentially (von der Embse, Jester, Roy, & Post, 2018). A caveat is that any discussion regarding this participant’s informal report is entirely observational. However, it is worth noting because it appears to be consistent with the idea that there is a potential for words to have varied meaning across individuals.

Despite the lack of the expected difference in memory for silently read verbs and nouns, that the production effect appears to generalise beyond nouns supports the claim (Bodner & MacLeod, 2016; Pals et al., 2018) that the effect is a viable learning strategy for applied settings.

**Hypothesis Three**

Lastly, the results do not support my hypothesis that the distinctiveness mechanism of the production effect is differentially affected by word-class. When people speak verbs there is a 24% increase in the proportion of words that they identify as ‘know’. When speaking nouns there is a 18% increase. The Bayes factor represents anecdotal evidence in favour of the null hypothesis that there is no difference in the size of the increase in ‘know’ judgments following production of verbs versus nouns. The pattern of explicit memory judgments in both groups appears to be consistent with the literature in that production benefits memory by improving people’s ability to form and access episodic memories (MacLeod et al., 2010).

In addition, there was no difference in the strength mechanism following production of verbs versus nouns. The Bayes factor represents anecdotal evidence in favour of the null hypothesis that there is no difference between word-classes in the size of the increase in ‘familiar’ judgements following production. For both word-classes there is a statistically non-significant 12% increase in the proportion of total words that they identify as ‘familiar’. This is consistent with Fawcett’s (2013) claim that the small implicit memory improvement of production is not consistently observed in experimental settings, rather it is detectable mostly at the level of a meta-analysis.

Regarding the effect of the independent ‘familiar/know’ adjustment, the results appear to support Fawcett and Ozubko’s (2016) claim that unadjusted ‘familiar’ values underestimate the memory benefits of the strength mechanism on item-recognition. An analysis of people’s unadjusted ‘familiar’ judgment proportions would suggest that people’s ‘familiar’ judgments decrease following production of verbs. This would indicate that production results in a weaker implicit memory for spoken verbs than memory for spoken nouns.

However, this interpretation would conflict with what is known about the production effect (Fawcett, 2013). The effect of the strength mechanism is small, but it is detectable at the level of meta-analysis. Furthermore, it does not appear to be logical that production of a word would result in a weaker representation of the word in implicit memory, regardless of word-class. Therefore, because it appears to be in contrast with the existing literature and not logical, it is probable that the independent adjustment allows for a more accurate interpretation of the data.

## Limitations

Regarding potential limitations of my design, using a backwards-counting filler task may not have been optimal. It is conceivable that there may have been a differential effect between individuals of mathematics-induced anxiety on test performance (Chang, & Beilock, 2016). It is unlikely that each participant was equally anxious so the observed variation in people's performance might partly have been due to anxiety, not the effect of production using different classes of words (von der Embse et al., 2018). However, the size of the observed production effect was large, which argues against a confounding effect.

Furthermore, it is a possibility that the study and test lists may not have contained enough words. There were clearly enough words to observe a production effect which was the main aim of my research, but there may not have been enough to observe a difference in 'familiar/know' judgments. The independent adjustment did not leave many opportunities for 'familiar' responses for some individuals, which may have potentially masked a difference between word-classes. However, a decision was made not to place excessive memory demands on participants, and the size of the lists in my design were consistent with past research (Bodner & Taikh, 2012; Jones & Pyc, 2014; MacLeod et al., 2010). Future research could follow that increases the number of words on the study list, possibly by combining verbs and nouns into a single list as I have evidenced that no difference of word-class is likely.

Lastly, my experiment may have been underpowered in terms of finding verb-noun differences in the size of the production effect. While the available evidence favours the null, it is only anecdotal. This has potentially resulted from low participant numbers creating a relatively large error variance. However, there is moderate evidence to suggest that there are no memory differences between silently read word-classes, which would argue against a differential production effect for verbs and nouns. However, this evidence is only moderate, so definitive statements cannot be made. Future researchers may be better able to find more substantial evidence if more participants are recruited.

**Future Directions**

Intransitive verbs (that do not require a direct object) such as 'relax' and 'breathe' appear to be notable for participants. Future research could be conducted that only uses transitive verbs (that require a direct object) to exaggerate any potential contextual differences between word-classes. I chose not to do so because the overall aim of my research was to examine the suitability of the production effect for real-world learning environments. Dividing verbs up into further categories would have detrimental to that goal because both transitive and intransitive verbs are commonly used.

**Conclusion**

In conclusion, the production effect is a large effect that reliably improves people's memory for spoken words. Furthermore, the size of the memory improvement appears not to differ between word-classes. Consequently, people's memory judgments following production also appear not to differ between word-classes. Within the constraints of low participant numbers, these results potentially support the idea that the production effect could be successfully used to improve both teaching and learning strategies (Pals et al., 2018) because it is not restricted to a memory benefit only for nouns.

## References

Bodner, G. E., & MacLeod, C. M. (2016). The benefits of studying by production ... and of studying production: Introduction to the special issue on the production effect in memory. *Canadian Journal of Experimental Psychology*, *70*(2), 89–92. doi:10.1037/cep0000094

Bodner, G. E., & Taikh, A. (2012). Reassessing the basis of the production effect in memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *38*(6), 1711–1719. doi:10.1037/a0028466

Chang, H., & Beilock, S. L. (2016). The math anxiety-math performance link and its relation to individual and environmental factors: A review of current behavioral and psychophysiological research. *Current Opinion in Behavioral Sciences*, *10*, 33–38. doi:10.1016/j.cobeha.2016.04.011

Dewhurst, S. A. Rackie, J. M. & van Esch, L. (2016). Not lost in translation: Writing auditorily presented words at study increases correct recognition “at no cost”. *Journal of Cognitive Psychology*, *28*(4), 437–442. doi:10.1080/20445911.2016 .1145684

Dodson, C. S., & Schacter, D. L. (2001). “If I had said it I would have remembered it”: Reducing false memories with a distinctiveness heuristic. *Psychonomic Bulletin & Review*, *8*(1), 155-161. doi:10.3758/BF03196152

Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, *1*(2), 170–177. doi:10.1037/1082-989X.1.2.170

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. doi:10.3758/BF03193146

Fawcett, J. M. (2013). The production effect benefits performance in between-subject designs: A meta-analysis. *Acta Psychologica*, *142*(1), 1–5. doi:10.1016

/j.actpsy.2012.10.001

Fawcett, J. M., & Ozubko, J. D. (2016). Familiarity, but not recollection, supports the between-subject production effect in recognition memory. *Canadian Journal of Experimental Psychology*, *70*(2), 99–115. doi:10.1037/cep0000089

Forrin, N. D., Jonker, T. R., & MacLeod, C. M. (2014). Production improves memory equivalently following elaborative vs non-elaborative processing. *Memory*, *22*(5), 470-480. doi:10.1080/09658211.2013.798417

Forrin, N. D., & MacLeod, C. M. (2017). This time it's personal: The memory benefit of hearing oneself. *Memory*, *26*(4), 574. doi:10.1080/09658211.2017 .1383434

Forrin, N. D., & MacLeod, C. M. (2018). Cross-modality translations improve recognition by reducing false alarms. *Memory*, *26*(1), 53-58. doi:10.1080 /09658211.2017.1321129

Forrin, N. D., MacLeod, C. M., & Ozubko, J. D. (2012). Widening the boundaries of the production effect. *Memory & Cognition*, *40*(7), 1046–1055. doi:10.3758 /s13421-012-0210-8

Gathercole, S. E., & Conway, M. A. (1988). Exploring long-term modality effects: Vocalization leads to best retention. *Memory & Cognition, 16*(2), 110–119. doi:10.3758/BF03213478

Han, S. H., & Kim, M. S. (2004). Visual search does not remain efficient when executive working memory is working. *Psychological Science*, *15*(9), 623–628. doi:10.1111/j.0956-7976.2004.00730.x

Hopkins, R. H., & Edwards, R. E. (1972). Pronunciation effects in recognition memory. *Journal of Verbal Learning and Verbal Behavior*, *11*(4), 534–537. doi:10.1016/S0022-5371(72)80036-7

Huttenlocher, J., & Lui, F. (1979). The semantic organization of some simple nouns and verbs. *Journal of Verbal Learning and Verbal Behavior*, *18*(2), 141–162.

doi:10.1016/S0022-5371(79)90091-4

Icht, M., Bergerzon-Biton, O., & Mama, Y. (2016). The production effect in adults with dysarthria: Improving long-term verbal memory by vocal production, *Neuropsychological Rehabilitation*, 1–13. doi:10.1080/09602011.2016.1272466

Icht, M., & Mama, Y. (2015). The production effect in memory: A prominent mnemonic in children. *Journal of Child Language*, *42*(5), 1102–1124. doi:10.1017/S0305000914000713

Jones, A. C., & Pyc, M. A. (2014). The production effect: Costs and benefits in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(1), 300–305. doi:10.1037/a0033337

Kersten, A. W., & Earles, J. L. (2004). Semantic context influences memory for verbs more than memory for nouns. *Memory & Cognition*, *32*(2), 198-211. doi: 10.3758/BF03196852

Lakoff, G., & Johnson, M. (2003). *Metaphors We Live By*. London, England: The University of Chicago Press.

Lin, O. Y. H. & MacLeod, C. M. (2012). Aging and the production effect: A test of the distinctiveness account. *Canadian Journal of Experimental Psychology*, *66*(3), 212–216. doi:10.1037/a0028309

MacLeod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(3), 671–685. doi:10.1037 /a0018785

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris*, *102*(1–3), 59–70. doi:10.1016/j.jphysparis.2008.03.004

Mama, Y., & Icht, M. (2016). Auditioning the distinctiveness account: Expanding the production effect to the auditory modality reveals the superiority of writing over vocalising. *Memory, 24*(1), 98–113. doi:10.1080/09658211.2014.986135

Mama, Y., & Icht, M. (2018). Production on hold: Delaying vocal production enhances the production effect in free recall. *Memory, 26*(5), 589-602. doi:10.1080/09658211.2017.1384496

Ozubko, J. D., & MacLeod, C. M. (2010). The production effect in memory: Evidence that distinctiveness underlies the benefit. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(6), 1543–1547. doi:10.1037 /a0020604

Pals, F. F., Tolboom, J. L., Suhre, C. J., & van Geert, P. L. (2018). Memorisation methods in science education: Tactics to improve the teaching and learning practice. *International Journal of Science Education*, *40*(2), 227–241. doi:10.1080 /09500693.2017.1407885

Quinlan, C. K., & Taylor, T. L. (2013). Enhancing the production effect in memory. *Memory, 21*(8), 904–915. doi:10.1080/09658211.2013.766754

Rackie, J. M., Brandt, K. R., & Eysenck, M. W. (2015). Interaction between mode of learning and subjective experience: Translation effects in long-term memory. *Memory, 23*(3), 318–328. doi:10.1080/09658211.2014.886701

Schubert, T. W., & Koole, S. L. (2009). The embodied self: Making a fist enhances men's power-related self-conceptions. *Journal of Experimental Social Psychology*, *45*(4), 828–834. doi:10.1016/j.jesp.2009.02.003

von der Embse, N., Jester, D., Roy, D., & Post, J. (2018). Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *Journal of Affective Disorders*, *227*, 483–493. doi:10.1016/j.jad.2017.11.048

Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., & Verhagen, J., . . . Meerhoff, F. (2016). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, *25*, 58–76. doi:10.3758 /s13423-017-1323-7

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625–636. doi:10.3758/BF03196322

# Appendix A

## Verb Item Pool

Words in blue font were spoken, words in black font were read silently.

| | | | |
|---|---|---|---|
| Attend | Describe | Wriggle | Migrate |
| Hum | Connect | Add | Scrape |
| Close | Delay | Improve | Hug |
| Stare | Climb | Nod | Mend |
| Kiss | Pull | Detect | Hold |
| Scrub | Pop | Deliver | Acquire |
| Breathe | Juggle | Cough | Injure |
| Carry | Wash | Borrow | Admire |
| Disagree | Knit | Inform | Decorate |
| Press | Behave | Relax | Bake |
| Inject | Enjoy | Allow | Bury |
| Push | Drown | Blink | Agree |
| Pray | Destroy | Drag | Sniff |
| Fetch | Shrug | Decide | Stir |
| Extend | Achieve | Boil | Smash |
| Clap | Slap | Kick | Bounce |
| Unlock | Eat | Hang | Fasten |
| Discover | Squeeze | Lick | Scare |
| Stop | Arrange | Greet | Yawn |
| Beg | Knock | Accept | Kill |
| Disarm | Develop | Surprise | Tremble |
| Tickle | Strum | Educate | Disappear |
| Belong | Sigh | Imply | Check |
| Argue | Give | Clean | Rinse |
| Pour | Count | Shave | Curl |

**Noun Item Pool**

Words in blue font were spoken, words in black font were read silently.

| | | | |
|---|---|---|---|
| Exam | Road | Turnip | Gravity |
| Decision | Oven | Election | Sailor |
| Clothes | Camera | Village | Capital |
| Disease | Lesson | Daughter | Invention |
| Computer | Year | Kingdom | Pebble |
| Ocean | Apartment | Uncle | Beauty |
| Music | Neighbour | Uniform | Foundation |
| Porch | Queen | Attention | Furniture |
| Amount | History | Factory | Army |
| Basket | Medicine | Ladder | Kitchen |
| World | President | Peace | Century |
| Wheat | Nephew | Stranger | Education |
| Data | Kettle | Speech | Family |
| Avenue | Country | Wood | Invitation |
| Television | Victory | Summer | Afternoon |
| Dinner | River | Industry | Department |
| Bird | Language | Justice | Minute |
| Soup | Guardian | Wagon | Internet |
| Audience | Lake | Teacher | System |
| Child | Battery | Dirt | Attitude |
| Event | Direction | Theatre | Castle |
| Article | Entrance | Trousers | Shirt |
| Story | Meadow | Office | Orchard |
| Month | Island | Valley | Merchant |
| Forest | Speaker | Evening | Engine |

# Appendix B

## Ethics Approval Letter/ Information Sheet/ Consent Form/ Debrief Statement

Social Science Ethics Officer
Private Bag 01 Hobart
Tasmania 7001 Australia
Tel: (03) 6226 2763
Fax: (03) 6226 7148
Katherine.Shaw@utas.edu.au

HUMAN RESEARCH ETHICS COMMITTEE (TASMANIA) NETWORK

16 August 2018

Dr Michael Garry
Psychology
Private Bag 30

Dear Dr Garry

Re: MINIMAL RISK ETHICS APPLICATION APPROVAL
Ethics Ref: H0017563 - Saying and Doing: Verbs and the Production Effect

We are pleased to advise that acting on a mandate from the Tasmania Social Sciences HREC, the Chair of the committee considered and approved the above project on 16 August 2018.

This approval constitutes ethical clearance by the Tasmania Social Sciences Human Research Ethics Committee. The decision and authority to commence the associated research may be dependent on factors beyond the remit of the ethics review process. For example, your research may need ethics clearance from other organisations or review by your research governance coordinator or Head of Department. It is your responsibility to find out if the approval of other bodies or authorities is required. It is recommended that the proposed research should not commence until you have satisfied these requirements.

Please note that this approval is for four years and is conditional upon receipt of an annual Progress Report. Ethics approval for this project will lapse if a Progress Report is not submitted.

The following conditions apply to this approval. Failure to abide by these conditions may result in suspension or discontinuation of approval.

1. It is the responsibility of the Chief Investigator to ensure that all investigators are aware of the terms of approval, to ensure the project is conducted as approved by the Ethics Committee, and to notify the Committee if any investigators are added to, or cease involvement with, the project.

2. Complaints: If any complaints are received or ethical issues arise during the course of the project, investigators should advise the Executive Officer of the Ethics Committee on 03 6226 7479 or human.ethics@utas.edu.au.

3. Incidents or adverse effects: Investigators should notify the Ethics Committee immediately of any serious or unexpected adverse effects on participants or unforeseen events affecting the ethical acceptability of the project.

4. Amendments to Project: Modifications to the project must not proceed until approval is obtained from the Ethics Committee. Please submit an Amendment Form (available on our website) to notify the Ethics Committee of the proposed modifications.

5. Annual Report: Continued approval for this project is dependent on the submission of a Progress Report by the anniversary date of your approval. You will be sent a courtesy reminder closer to this date. **Failure to submit a Progress Report will mean that ethics approval for this project will lapse.**

6. Final Report: A Final Report and a copy of any published material arising from the project, either in full or abstract, must be provided at the end of the project.

Yours sincerely

Jude Vienna-Hallam
Acting Executive Officer – Social Science
Tasmania Social Sciences HREC

# Saying and Doing: Verbs and the Production Effect

For participants

## 1. Invitation

You are invited to participate in a study investigating the production effect and its effect on memory. This study is part of psychology Honours research project of Hayley Blowfield, Michael O'Leary and Zac Bain-Williams. The study is supervised by Dr Michael Garry from the Division of Psychology at the University of Tasmania.

## 2. What is the purpose of this study?

The purpose of this study is to investigate whether the way people study words affects how well the studied words are remembered. The results from this study may help teachers design activities to make learning more effective. We can't provide too many details at the start of the study as this could affect the results. Once you have completed your study session we will fully explain the aims and hypotheses.

## 3. Why have I been invited to participate?

You have been invited to participate because you are between the age of 18 and 50, speak English as my first language, have normal or corrected to normal vision, are not colour blind, and don't have a reading disorder, or neuromuscular or neurological disorder.

## 4. What will I be asked to do?

The study will involve a single session of up to 90 minutes duration. The session will be in either the Psychology Research Centre, or the Social Sciences Building on the Sandy Bay campus at the University of Tasmania.

At beginning of the session you will be asked to provide some information about yourself, such as your age and sex, and will be asked to complete a questionnaire to determine your hand preference. The experimenter will then assign you to a group which will determine what you will be asked to do in the remainder of the session. During the main part of the session you will be seated at a desk in front of a computer screen. The screen will present individual words, one after the other, several seconds apart. Some of the words will be written in white letters and other words will be written in blue letters. Approximately 80-100 words will be presented.

Regardless of which group you are in you will be asked to silently read all words written in WHITE letters. For words written in BLUE letters, what you will be asked to do will depend on which group you are in. You may be asked to say the blue words aloud, to perform a movement with your hands to mime the blue word, or to both say the word and perform the miming hand movement simultaneously. The experimenter will explain the precise instructions to you.

To allow us to confirm the accuracy of your vocal responses we will use a microphone to record your spoken responses. The recordings will be used to confirm that your response and how long it took you to respond. The recordings will be deleted once this data has been extracted.

5. Are there any possible benefits from participation in this study?

Your participation in this study will contribute to the scientific literature about a phenomenon called the production effect. This effect has been used to improve how students learn new information. This study will add to that body of knowledge.

Your participation in this study will provide valuable research training for three Psychology Honours students: Hayley Blowfield, Michael O'Leary and Zac Bain-Williams.

If you are a first year Psychology student you will receive up one hour course credit. If you are not a first-year psychology student, or aren't seeking course credit your name will be entered into a draw to win one of six $100 Coles-Myer vouchers.

6. Are there any possible risks from participation in this study?

There are no risks associated with participating in this study. If you become tired during the study, you can take an extended break between trials.

7. What if I change my mind during or after the study?

Participation in this study is entirely voluntary. You are welcome to withdraw from the study without consequence at any time if you wish. You are encouraged to do so if you are experiencing distress. If your data has already been collected, you can request to have it removed and destroyed up to 14 days after the date of your participation. It will not be possible to remove your data after this time as it will be de-identified for data analysis. However, once the data has been analysed and the thesis written your data will not be able to be removed.

8. What will happen to the information when this study is over?

All data will be replaced by an alphanumeric code, which will be recorded on your consent form. This will allow your data to be re-identified if necessary. Your identity will remain confidential for the purposes of the research and your information will not be provided to any sources without your knowledge. Any paper information will be stored of securely, in a locked cabinet at the University of Tasmania. Your electronic data will be stored securely on a password-protected computer at the University of Tasmania. Data will be kept for a minimum of 5 years after the date of publication.

9. How will the results of the study be published?

The final results will be reported an Honours thesis, which will be available to access through the UTAS Psychology Test Library. A brief summary of the thesis results will be available from 1st December 2018. Please contact the Chief Investigator, Dr Mike Garry (Michael.garry@utas.edu.au) if you would like to receive a copy of the summary.

10. What if I have questions about this study?

If you have any questions or concerns about the research, you are advised and encouraged to consult one of the student investigators (Hayley Blowfield, hayleyb4@utas.edu.au; Michael O'Leary, olearym@utas.edu.au; Zac Bain-Williams, zbain@utas.edu.au) or Mike Garry (chief investigator) on michael.garry@utas.edu.au at any time.

This study has been approved by the Tasmanian Social Sciences Human Research Ethics Committee. If you have concerns or complaints about the conduct of this study, please contact the Executive Officer of the HREC (Tasmania) Network on +61 3 6226 6254 or email human.ethics@utas.edu.au. The Executive Officer is the person nominated to receive complaints from research participants. Please quote ethics reference number: H0017563.

**Thank you for considering participating in our study.**

**You are welcome to keep this information sheet and refer back to it.**

**If you wish to participate in this study, please sign the attached consent form.**

UNIVERSITY *of* TASMANIA

*University of Tasmania*

**COLLEGE OF HEALTH and MEDICINE**

**School of Medicine**

*Participant Consent Form [Version 1] [16/07/2018]*

# Saying and Doing: Verbs and the production effect.

<u>This form is for participants</u>

1. I agree to take part in the research study named above.
2. I have read and understood the Information Sheet for this study.
3. The nature and possible effects of the study have been explained to me.
4. I understand that the study involves reading words presented on a computer monitor and that the study will take up to 90 minutes to complete.
5. I understand that a microphone will be used to record my voice during trials. This is to allow the researchers to confirm my responses have been accurately recorded. I also understand that these recordings will be erased once the accuracy of responses has been confirmed.
6. I understand that there are no risks associated with participating in this study
7. I understand that all research data will be securely stored on the University of Tasmania premises for a minimum of five years from the publication of the study results, and will then be destroyed.
8. Any questions that I have asked have been answered to my satisfaction.
9. I understand that the researcher(s) will maintain confidentiality and that any information I supply to the researcher(s) will be used only for the purposes of the research.
10. I understand that the results of the study will be published so that I cannot be identified as a participant.
11. I understand that my participation is voluntary and that I may withdraw at any time without any effect. If I so wish, I may request that any data I have supplied be withdrawn from the research , though this will only be possible up to 14 days following participation.

*Participant Consent Form [Version 1] [16/07/2018]*

Participant's name:

______________________________________________

Participant's signature:

___________________________________________

Date: ____________________

**Statement by Investigator**

☐ I have explained the project and the implications of participation in it to this volunteer and I believe that the consent is informed and that he/she understands the implications of participation.

If the Investigator has not had an opportunity to talk to participants prior to them participating, the following must be ticked.

☐ The participant has received the Information Sheet where my details have been provided so participants have had the opportunity to contact me prior to consenting to participate in this project.

Investigator's name:

______________________________________________

Investigator's signature:

___________________________________________

Date: ____________________

# Production effect participant debrief statement.

Now that your data collection session is complete I would like to explain the aims and hypotheses. The aim of this study is to investigate a research finding called the *production effect*. The production effect is the finding that when we try to remember a list of words, our memory is usually better for words that we have spoken out loud compared to words that we have only read silently.

In the study you just completed you spoke out loud, and/or mimed, some of the words that were presented, and silently read other words. Based on previous research on the production effect I expect to find that you will remember more words that you spoke out loud, or mimed, compared to words that you only read silently.

Another aim of this study is to investigate whether the production effect is found when the words studied are verbs (words that describe actions such as 'pinch', and 'hit'). This differs from previous studies which have used noun words (words that describe things such as apple, and car). One of the reasons I used verbs instead of nouns is because I want to know whether the production effect would also occur when people mime an action (e.g., mime the word 'hit'), not just when they speak a word out loud. Based on several lines of research I expect to find a production effect for mimed words that is as strong, or stronger in comparison with spoken words.

The reason I didn't tell you these hypotheses at the start of the study is because I didn't want to influence your memory strategies. To be able to accurately interpret the results I need to be sure that you were trying to remember all the words presented, regardless of whether you spoke, mimed or read them silently.

Now that you know what the hypotheses of the study are, I would like to ask whether you still consent your data being included in the study. As discussed at the start of the study you are free to withdraw your data if you choose. Do you have any questions?

[Experimenter will answer any questions.

If the participant requests to withdraw their data the experimenter will delete the data from the computer and destroy the signed consent form by tearing it up.

Once all the participant's questions have been answered and inclusion of data confirmed (or withdrawn), the experimenter will thank the participant for their participation and end the study session.]